



별첨 사본은 아래 출원의 원본과 동일함을 증명함.

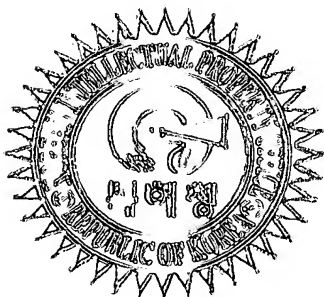
This is to certify that the following application annexed hereto  
is a true copy from the records of the Korean Intellectual  
Property Office.

165785

출원번호 : 10-2003-0004508  
Application Number

출원년월일 : 2003년 01월 23일  
Date of Application JAN 23, 2003

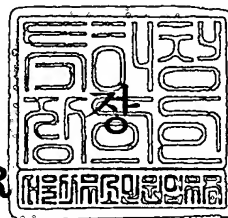
출원인 : 삼성전자주식회사  
Applicant(s) SAMSUNG ELECTRONICS CO., LTD.



2003      년      08      월      09      일

특      허      청

COMMISSIONER



## 【서지사항】

【서류명】	특허출원서
【권리구분】	특허
【수신처】	특허청장
【참조번호】	0005
【제출일자】	2003.01.23
【국제특허분류】	H04N
【발명의 명칭】	비디오 코딩시의 비트율 제어 방법 및 장치
【발명의 영문명칭】	Rate control method and apparatus in video coding
【출원인】	
【명칭】	삼성전자 주식회사
【출원인코드】	1-1998-104271-3
【대리인】	
【성명】	이영필
【대리인코드】	9-1998-000334-6
【포괄위임등록번호】	2003-003435-0
【대리인】	
【성명】	이해영
【대리인코드】	9-1999-000227-4
【포괄위임등록번호】	2003-003436-7
【발명자】	
【성명의 국문표기】	장승기
【성명의 영문표기】	CHANG, Seung Gi
【주민등록번호】	660810-1816817
【우편번호】	137-843
【주소】	서울특별시 서초구 방배1동 928-31 세종아트빌 202호
【국적】	KR
【발명자】	
【성명의 국문표기】	장지밍
【성명의 영문표기】	ZHANG, Zhiming
【주소】	경기도 수원시 팔달구 영통동 동보신명아파트 202동 1306호
【국적】	CN

【공개형태】	간행물 발표
【공개일자】	2002.09.03
【심사청구】	청구
【취지】	특허법 제42조의 규정에 의한 출원, 특허법 제60조의 규정에 의한 출원심사를 청구합니다. 대리인 이영필 (인) 대리인 이해영 (인)
【수수료】	
【기본출원료】	20 면 29,000 원
【가산출원료】	8 면 8,000 원
【우선권주장료】	0 건 0 원
【심사청구료】	9 항 397,000 원
【합계】	434,000 원
【첨부서류】	1. 요약서·명세서(도면)_1통 2. 공지예외적용대상(신규성상실의예외, 출원시의특례)규정을 적용받기 위한 증명서류_1통

## 【요약서】

## 【요약】

본 발명은 엠펙-4(MPEG-4) 비디오 코딩에 관한 것으로, 구체적으로는 엠펙-4 비디오 코딩에서의 비트율 제어(rate control) 방법 및 장치에 관한 것이다. 본 발명의 비트율 제어방법은 비디오 프레임을 입력받고, 디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 비트율 왜곡(RD) 버퍼에 메모리를 할당하는 초기화 수행 단계; 입력받은 비디오 프레임에 대하여 인트라 코딩을 수행하는 단계; RD 버퍼를 업데이트하는 단계; RD 버퍼가 언더플로우나 오버플로우 되지 않도록 다음 입력받은 프레임을 스킵하는 포스트 프레임 스킵 수행 단계; 다음 비디오 프레임을 입력받아 첫번째 비디오 오브젝트 플레인 인가의 여부에 따라서 양자화값을 추정하거나 인터 코딩을 수행하는 단계; 인터 코딩을 수행한 후에 입력받은 프레임의 수에 따라서 역전파(back propagation) 모델 업데이트를 수행하거나 자가 구축 제어를 수행한 후 포스트 프레임 스킵을 수행하는 단계; 및 다음 비디오 프레임을 입력받아 모든 프레임에 대하여 상술한 과정을 다시 수행하는 단계를 구비한다. 본 발명은 RD 모델을 자가구축하기 때문에 복수의 샘플, 복수의 양자화기 및 선행 트레이닝과정이 필요하지 않아 비트율 제어에 걸리는 시간이 많이 단축되는 효과가 있다.

## 【대표도】

도 5

**【명세서】****【발명의 명칭】**

비디오 코딩시의 비트율 제어 방법 및 장치{Rate control method and apparatus in video coding}

**【도면의 간단한 설명】**

도 1은 MPEG-4의 계층구조를 도시한 도면이다.

도 2는 LPN 의 기본적인 학습 구조를 도시한 도면이다.

도 3은 표준 SOM(Self Organizing Mapping) 알고리즘을 사용하여 RD 모델을 자가 구축하는 방법의 흐름도이다.

도 4는 SOLPN의 기본구조를 이용하여 시스템을 자가구축하는 과정을 설명하는 순서도이다.

도 5는 본 발명의 비트율 제어방법의 상세 흐름도이다.

도 6은 본 발명의 비트율 제어장치의 블록도이다.

**【발명의 상세한 설명】****【발명의 목적】****【발명이 속하는 기술분야 및 그 분야의 종래기술】**

<7> 본 발명은 엠펙-4(MPEG-4) 비디오 코딩에 관한 것으로, 구체적으로는 엠펙-4 비디오 코딩에서의 비트율 제어(rate control) 방법 및 장치에 관한 것이다.

<8> 엠펙(Moving Picture Expert Group : MPEG)에서 제안하고 있는 동영상 압축방법은 시간적 중복(temporal redundancy) 및 공간적 중복(spatial redundancy)을 제거하는 방

법에 기반을 두고 있다. 즉, 이동보상 압축방법을 이용하여 시간적 중복을 제거하고, 정지화상에 DCT(Discrete Cosine Transform)를 적용하여 공간적 중복을 제거한다.

- <9> 이들 MPEG 과 관련된 압축 표준들 중에서 MPEG-4 는 객체(object) 기반의 동영상 압축기술이다. MPEG-4 가 기존의 표준방식과 다른 점은 임의 형상의 물체와 객체를 개별적으로 부호화한다는 것이다.
- <10> 도 1은 MPEG-4의 계층구조를 도시한 도면이다.
- <11> 화상의 전체적인 시퀀스를 비디오 세션(Video Session : VS)(110)라고 하는데, 비디오 세션(VS)(110)에는 1개 이상의 비디오 오브젝트(Video Object : VO)(120)가 존재한다. 예를 들어 배경 가운데에 인물이 존재하는 경우 인물만의 연속 움직임을 하나의 비디오 오브젝트(VO)로 기술할 수도 있고, 또 배경 시퀀스(sequence)도 별개로 기술할 수 있다. 각 비디오 오브젝트(VO)(120) 밑에는 비디오 오브젝트 계층(Video Object Layer : VOL)(130)이라는 계층이 있다. 비디오 오브젝트 계층(VOL)(130)은 각각의 비디오 오브젝트(VO)에 대해 복수의 공간 해상도 또는 시간 해상도를 부여한다.
- <12> 최하위 층에 있는 비디오 오브젝트 평면(Video Object Plane : VOP)(150)은 각 비디오 오브젝트(VO)의 해상도에 대해 '어느 순간'의 데이터를 의미한다. 또한 비디오 오브젝트 계층(VOL)(130)과 비디오 오브젝트 평면(VOP)(150) 사이에는 VOP의 그룹(Group of VOP)(140)라는 새로운 계층이 존재한다. 이는 랜덤 액세스를 수행하기 위한 것으로 이 계층이 존재하는 경우 시간 방향의 예측을 하지 않는 모드로부터 부호화가 시작된다.
- <13> 객체기반 부호화(Object based coding) 방법인 MPEG-4 에서 가변 비트율(Variable Bit Rate : VBR) 서비스의 제공시에 고정된 양자화 비트를 사용하는 방법이 여전히 많이

사용되고 있다. 이렇게 고정된 양자화 비트를 사용하면 출력되는 비트 스트림의 데이터 양이 너무 심하게 변화하면 출력버퍼의 오버플로우(overflow)나 언더플로우(underflow) 현상이 일어날 가능성이 많아진다. 또한, 항등 비트율(Constant Bit Rate : CBR) 서비스에서는 양자화 비트를 유연하게 조정하여, 출력되는 데이터의 양이 입력된 영상의 종류에 관계없이 일정하도록 조절할 필요가 있는데, 이를 비트율 제어(rate control)라고 한다.

<14> 비트율 제어방법에는 출력되는 데이터의 양을 프레임 단위로 조절하는 프레임 기반의 비트율 제어방법과 출력되는 데이터의 양을 프레임 내의 각 매크로 블록단위로 조절하는 매크로 블록 기반의 비트율 제어 방법이 존재한다. 매크로 블록 기반의 비트율 제어방법이 더 정확하게 비트율을 제어할 수 있지만, 제어방법이 너무 복잡하고 어려워 프레임 기반의 비트율 제어방법이 일반적으로 많이 사용된다.

<15> 실시간 비디오 통신에 있어서 비트율 제어(rate control)는 비디오 인코더에서 매우 중요하다. 비트율 제어에 있어서 중요하게 고려할 점은 단대단 지연(end-to-end delay) 조건을 충족시키고, 인코딩시 사용되는 버퍼에 오버플로우(overflow)나 언더플로우(underflow)가 발생되지 않도록 비디오 인코더의 비트율 왜곡(RD) 기능을 추정하는 것이다. 버퍼에 있는 비트 수가 너무 크면 인코더는 버퍼의 지연을 줄이고 버퍼의 오버플로우를 방지하기 위하여 인코딩 프레임을 스킵(skip)한다. 오버플로우 등으로 인하여 프레임이 스킵되면 인코딩된 비디오 시퀀스에 있어서 불연속 때문에 디코딩된 영상의 움직임이 부자연스러워지는 결과를 초래한다.

<16> 프레임 기반의 비디오를 인코딩할 때, 비트의 한계를 고려하여 적절한 양자화기를 선택하여야 한다. 이때 적절한, 그리고 적응적인(adaptive) 비트율 왜곡(Rate

Distortion : RD) 모델 구축시 이러한 점을 반드시 고려하여야 한다. 한편 SOLPN(Self Organizing Learning Petri Net) 이 있다.

- <17>      페트리넷(Petri Net)은 1960년대 서독의 칼 페트리(Carl Petri)에 의해서 제안된 것으로 다양한 상황을 모형화 할 수 있는 것으로 플레이스(place)와 트랜지션(transition)이라고 하는 두 종류의 노드(node)로 구성되며, 아크(arc)를 통해 서로 다른 타입의 노드와 연결된다. 여기서, 트랜지션은 입력신호에 대한 출력신호를 발생시키는 함수를 말하며, 플레이스는 어떤 입출력 신호를 저장하는 공간이라고 할 수 있다. 이러한 페트리넷에 신경망과 같은 학습능력을 부가한 것이 LPN(Learning Petri Net)이다.
- <18>      도 2는 LPN 의 기본적인 학습 구조를 도시한 도면이다.
- <19>      도면에서 입력 트랜지션을 제외한 각 트랜지션은 소정 개수의 입력 플레이스와 출력 플레이스를 가진다. 그리고 서로 다른 트랜지션은 단순성을 위하여 입력 플레이스나 출력 플레이스를 동시에 갖지 않는 것으로 간주된다. 도 1에서는 트랜지션 및 플레이스의 개수를 간단하게 나타내었으나 실제로는 더 많은 트랜지션 및 플레이스들이 병렬 또는 직렬로 서로 연결되며, 더 다양한 형태를 가질 수 있다.
- <20>      상술한 LPN은 신경망의 학습 및 재현 능력을 가질 수 있다. 그러나 LPN은 분포함수의 특성을 가지기 때문에 신경망과는 다르며, 동시에 정상적인 신경망과 같이 전문가의 경험에 의해 LPN의 파라미터들이 미리 결정된다. 이러한 LPN은 입력 레이어와 출력 레이어 사이의 모든 트랜지션과 플레이스의 숫자 및 그 연결이 사용자의 경험에 의해 미리 고정되므로 출력값이 다소 정확하지 못하다는 문제점이 있어 SOLPN 이 제안되었다.



- <21> SOLPN 은 사용자의 경험에 의한 시스템에서 학습이 수행되는 것이 아니라 샘플들의 트레이닝을 통해 구축된 시스템을 통해 학습이 수행되므로 학습속도가 빠르고, 보다 정확한 모델링이 가능한 자가구축 LPN 을 말한다.
- <22> DCT 기반의 비디오 인코더에서는 여러가지 비트율 왜곡(RD) 모델이 사용된다. 그중 하나로 각각의 이미지 블록을 인코딩하고, 가장 좋은 파라미터를 지능적으로 선택하는 모델이 있다. 그러나, 이러한 모델을 사용한 방법은 계산이 너무 복잡하기 때문에 실시간 인코딩시에 적용하기에는 적합하지 않다.
- <23> 다른 방법으로 다른 수학적 모델에 기초하여 양자화기(quantizer)를 선택하고, 제어 파라미터(control parameter)가 코딩 시스템의 비트율 왜곡(RD) 데이터로부터 추정된다. 이러한 방법이 실시간 응용에는 적합하지만, 이 방법에서는 프레임의 스킵이 빈번하게 발생하고, 저지연(low-delay) 응용 서비스의 제공시에는 채널의 대역(bandwidth)을 낭비하게 된다. 또한 높은 코딩 효율을 얻기 위해서, 비트율 왜곡(RD) 알고리즘은 점점 더 복잡해지고, 실험제어 파라미터를 얻기 위해서 더 많은 실험이 행해질 필요가 있다.
- <24> Electronics Letter 지(vol 36. no 12 pp.1041-1158, 2000)에서는 "Rate control algorithm using SOFM-based neural network" 이라는 제목의 논문에서 비트율 제어 스킵(scheme)에 기반을 둔 자가 구축(self organizing) 맵이 개시되었다. 이것은 신경 분류자(neural classifier)에 의한 프레임 기반의 전역(global) 비트율 왜곡(RD) 모델을 구축한다. 개시된 맵을 사용하여 좋은 결과를 얻을 수 있지만, 오프라인(off-line)상의 트레이닝을 필요로 한다. 즉, 복수의 비디오 샘플을 가지고 이미지 특성에서 가능한 모든 변형을 수집하기 위하여 1에서 31까지 변화하는 고정된 양자화값(QP)을 사용하여야 한다. 이 방법은 온라인 제어에는 적합하지 않고 구성(configuration)을 업데이트하는

것도 용이하지 않으며 대응되는 신경망(neural network) 구조도 쉽게 업데이트할 수 없다는 문제점이 있다.

**【발명이 이루고자 하는 기술적 과제】**

<25> 본 발명이 이루고자 하는 기술적 과제는, 비디오의 동적 성질을 획득하고 비디오 코딩시에 적절한 양자화기를 적응적으로 얻는 SOLPN 을 이용한 비트율 제어 방법 및 장치를 제공하는데 있다.

**【발명의 구성 및 작용】**

<26> 상기의 과제를 이루기 위하여 본 발명에 의한 비트율 제어방법은, 비디오 프레임을 입력받고, 디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 비트율 왜곡(RD) 버퍼에 메모리를 할당하는 초기화 수행 단계; 상기 입력받은 비디오 프레임에 대하여 인터라 코딩을 수행하는 단계; 상기 RD 버퍼를 업데이트하는 단계; 상기 RD 버퍼가 언더플로우나 오버플로우 되지 않도록 다음 프레임의 인코딩 여부를 판단하여 다음 입력받은 프레임을 스킵하는 포스트 프레임 스킵 수행 단계; 다음 비디오 프레임을 입력받아 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가의 여부에 따라서 양자화값을 추정하거나 인터 코딩을 수행하는 단계; 상기 인터 코딩을 수행한 후에 상기 입력받은 프레임의 수에 따라서 역전파(back propagation) 모델 업데이트를 수행하거나 자가 구축 제어를 수행한 후에 상기 포스트 프레임 스킵 수행 단계를 다시 수행하는 단계; 및 다음 비디오 프레임을 입력받아 모든 프레임에 대하여 상기 양자화값 추정단계 내지 자가 구축 제어 수행 단계를 수행하는 단계를 구비한다.

- <27>      상기의 과제를 이루기 위하여 본 발명에 의한 비트율 제어장치는, 비디오 스트림을 입력받아 코딩에 필요한 버퍼를 초기화하는 프리 인코딩부; 상기 입력받은 비디오 스트림에 대하여 인터 코딩과 인트라 코딩을 수행하는 인코딩부; 상기 코딩된 비디오 데이터에 따라서 상기 버퍼를 업데이트하고, 프레임 스킵제어를 수행하여 비트율을 조절하는 포스트 인코딩부; 다음 프레임을 입력받는 타임 인스턴트 업데이트부; 및 상기 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가를 판단하는 판단부를 구비한다.
- <28>      상기한 과제를 이루기 위하여 본 발명에서는, 상기 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체를 제공한다.
- <29>      이하, 첨부된 도면을 참조하여 본 발명에 따른 바람직한 일실시예를 상세히 설명한다.
- <30>      도 3은 표준 SOM(Self Organizing Mapping) 알고리즘을 사용하여 RD 모델을 자가 구축하는 방법의 흐름도이다.
- <31>      표준 SOM 알고리즘은 퍼지 로직에 사용되는 전통적인 알고리즘으로 트레이닝 샘플을 이용하여 고유의 유사성을 찾고 네트워크를 자가구축하는 방식의 알고리즘을 말한다.
- <32>      우선 하나의 샘플 비디오 프레임을 입력받는다. 그리고 입력받은 비디오 프레임을 미리 설정한 양자화값(QP) 값을 사용하여 인코딩한다(S310). 그리고 난 후 표준 SOM 알고리즘을 사용하여 자가구축(self-organize)한다(S320). 자가구축하는 과정은 다음의 도 4를 참조하여 상세히 설명한다.
- <33>      도 4는 SOLPN의 기본구조를 이용하여 시스템을 자가구축하는 과정을 설명하는 순서도이다.

- <34> 시스템을 구축하기 위하여 먼저, 입력값 및 출력값을 미리 알고 있는 샘플셋 (sample set)으로 순차적으로 트레이닝한다(S410). 즉, 제1 번째 샘플의 트레이닝을 통해 제1 시스템 파라미터를 설정한 후(S420), 제1 시스템 파라미터에 의해 구축된 시스템에 제2 번째 샘플을 트레이닝한다(S430). 그리고 제2 번째 트레이닝 샘플의 트레이닝 결과에 따른 출력값과 미리 알고 있는 트레이닝 샘플의 출력값 사이의 오차값을 산출하여 임계값과 비교한다(S440). 이때, 제2 샘플 트레이닝 결과에 따른 출력값과 미리 알고 있는 트레이닝 샘플의 출력값 사이의 오차값이 소정의 임계값 이하이면, 제1 시스템에 제3 샘플을 연속적으로 트레이닝 한다. 만일 제2 샘플 트레이닝 결과에 따른 출력값과 미리 알고 있는 트레이닝 샘플의 출력값 사이의 오차값이 임계값보다 크면, 제2 샘플에 따른 제2 시스템 파라미터를 생성하여 이를 기존에 존재하는 제 1 시스템 파라미터에 더한다 (S450).
- <35> 이후, 순차적으로 입력되는 후속 샘플들에 대해서도 각각 트레이닝 결과에 따른 출력값의 오차값에 따라 새로운 시스템 파라미터를 생성하거나 이전 결정된 시스템 파라미터에 후속 샘플을 계속 트레이닝 한다. 그리고 최종 샘플의 트레이닝 완료 상태를 검출한 후(S460), 샘플들의 순차적 트레이닝을 소정회수 반복하여 시스템 자가구축과정을 마친다(S470).
- <36> 도 4를 참조하여 설명한 자가구축과정을 수행한 후에,  $QP > 31$  인가를 판단하여 (S330),  $QP > 31$ 이 아니면, QP를 하나 증가시키고(S340), 그렇지 않으면 복수의 비디오 샘플(multiple video sample)에 대하여 상술한 과정을 반복한다(S350).
- <37> 그리고, 상술한 과정에서 구축된 네트워크를 사용하여 비디오를 인코딩 한다 (S360). 우선 사용자가 미리 정의한 양자화기(user-predefined quantizer)를 사용하여

첫번째 I 프레임과 P 프레임을 인코딩한다. 그리고 나서 RD 버퍼를 업데이트하거나 처리할 프레임을 스킵할 것인가 아닌가를 제어하는 포스트 프레임 스킵 제어(post frame-skip control)를 수행한다.

<38> 신경망(neural network)을 이용하여 양자화값(QP)을 추정하고, 현재의 이미지 프레임을 인코딩한 후 버퍼 제어를 수행한다. 즉, 만일 버퍼의 상태가 80% 이상이면 표준 VM18.0에 기술된 바와 같이 다음 프레임은 생략(skipped)될 것이다. 다음 프레임을 입력받아 인코딩하거나 종료한다.

<39> 상술한 방법은 신경망을 복수번 트레이닝시키는데 복수의 비디오 샘플이 필요하고, 너무 많은 장치들(units)이 필요하며, 비디오의 처리(processing)나 인코딩시에 구축된 네트워크의 구조가 쉽게 변형(modify)될 수 없다는 문제점이 있다.

<40> 종래의 방법과 같이 네트워크를 미리 트레이닝시키는 것은 너무 많은 시간이 걸리고 샘플링 비디오를 선택하는 것이 실질적으로 매우 어렵다. 따라서 본 발명에서는 비트스트림(bitstream)을 인코딩할 때 네트워크를 자가 구축하는 방법을 제공한다. 따라서 미리 트레이닝하는 시간이 불필요하다. 그리고, 네트워크 모델이 인코더 그 자체에 의해서 구축되기 때문에, 비트스트림의 특성을 적응적으로 얻을 수 있다.

<41> 본 발명의 기본적인 개념은 다음과 같다. 즉, N 개의 비디오 프레임을 인코딩하기 전에, SOLPN 알고리즘을 사용한 RD 모델을 자가 구축한다. 자가 구축 과정이 완료된 후, 양자화기를 추정하기 위하여 상술한 자가 구축 네트워크를 사용한다. 그리고 추정 오차(estimation error)에 따라서 네트워크의 구조를 동적으로 변화시킨다.

<42> 도 5는 본 발명의 비트율 제어방법의 상세 흐름도이다.

- <43> 비디오 프레임이 순차적으로 하나씩 입력되었을 때 본 발명의 비트율 제어방법을 도 5를 참조하여 상세히 설명하면 다음과 같다.
- <44> (1) 초기화를 수행한다(S502). 초기화 단계에서는, RD 버퍼에 메모리를 할당하는 등의 기능을 수행한다.
- <45> (2) 인트라 코딩(Intra coding)을 수행한다(S504). 즉, 프레임 0에 대해서 DCT, 양자화 및 VLC(Variable Length Coding)을 포함하여 인트라 코딩을 수행한다.
- <46> (3) RD 버퍼를 업데이트한다(S506). RD 버퍼는 디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 버퍼를 말한다.
- <47> (4) 포스트 프레임 스킵(post-frame skip) 과정을 수행한다(S508). 즉, 인코딩된 비트 스트림에 따라서 RD 버퍼가 언더플로우나 오버플로우 되지 않도록 다음 프레임의 인코딩 여부를 판단한다.
- <48> (5) 타임 인스턴트 업데이트(time instant update)를 수행한다(S510). 즉, 다음 프레임을 선택하여 입력받는다.
- <49> (6) 선택하여 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가를 판단한다(S512). 만일 첫번째 비디오 오브젝트 플레인(VOP)이면 인터 코딩(Inter Coding)을 수행하고, 그렇지 않으면 타겟 추정(target estimation)을 수행한다. 현재 설명하고 있는 과정에서는 첫번째 비디오 오브젝트 플레인(VOP)에 해당하므로 인터 코딩을 수행한다.

- <50> (7) 인터 코딩(Inter Coding)을 수행한다(S514). 즉, 선택한 프레임 1에 대해서 DCT, 양자화(quantization) 및 가변장 길이 부호화(Variable Length Coding : VLC)을 포함하여 인터 코딩을 수행한다.
- <51> (8) 입력받은 프레임의 수가 미리 설정한 소정의 상수 N 보다 작은가를 판단한다(S516). 만일 소정의 상수 N 보다 작으면 자가구축 제어(self-organizing control)를 수행하고, 그렇지 않으면 역전파(Back Propagation : BP) 모델 업데이트 제어를 수행한다. 현재 설명하고 있는 과정에서 처리되고 있는 프레임은 프레임 1이므로 모두 2개의 프레임을 처리한 것이 되므로 자가 구축 제어 단계를 수행하게 된다.
- <52> (9) 자가 구축 제어를 수행한다. 자가 구축 제어는 코딩결과에 따라서 네트워크를 구축하고 RD 버퍼를 업데이트 하는 과정이다.
- <53> (10) 포스트 프레임 스킵을 수행한다(S508).
- <54> (11) 타임 인스턴트 업데이트를 수행한다(S510). 전단계의 결과에 따라서 인코딩될 다음 프레임 i 를 선택한다.
- <55> (12) 이번에 입력된 프레임은 첫번째 비디오 오브젝트 플레인(VOP)이 아니므로 타겟 추정(target estimation)을 수행한다(S520). 즉, RD 버퍼의 상태에 따라서, 어느 정도의 비트들이 프레임 i 에 할당되어야 하는가를 추정한다.
- <56> (13) 조인트 버퍼 제어(joint buffer control)를 수행한다(S522). 네트워크의 구성과 상기 타겟 추정결과에 따라서 RD 버퍼를 수정한다.
- <57> (14) 구축된 네트워크를 이용하여 양자화값(QP)을 추정한다(S524). 즉, 구축된 네트워크에 따라서, 인코딩시에 사용될 양자화값(QP)을 추정한다. 추정하는 과정은 다음과

같다. 우선, 입력 벡터(input vector)를 입력 레이어(input layer)에 둔다. 그리고 구축된 네트워크로부터 예측 출력(prediction output)을 얻는다. 그리고 타겟 출력(target output)을 양자화값(QP)  $\pm$  이내의 오차로 예측한다(일관된 화질을 보장하기 위하여, 인접 프레임간의 차이가  $\pm$ 를 넘지 않도록 제한한다). 그리고 winner를 찾는다. 즉 타겟 비트(target bit) 할당을 가진 가장 근접한 출력의 적절한 양자화값(QP)을 찾는다. 선택된 양자화값(QP)을 양자화기에 할당한다.

- <58>        (15) 인터코딩을 수행한다(S514).
- <59>        (16) 이번에 입력된 프레임도 N 보다 작으므로(S516), 자가 구축 단계를 수행한다.
- <60>        (17) 자가 구축과정을 수행한다(S518). 인코딩된 비트 스트림에 따라서 네트워크는 자가구축과정을 계속한다.
- <61>        (18) 포스트 프레임 스킵 제어를 수행한다(S508).
- <62>        (19) 타임 인스턴트 업데이트를 수행한다(S510). 다음 프레임을 입력받아, 프레임의 수가 N을 초과할 때까지 상술한 12-19 단계를 반복한다.
- <63>        (20) 타겟 추정을 수행한다(S520).
- <64>        (21) 조인트 버퍼 제어를 수행한다(S522).
- <65>        (22) 구축된 네트워크를 이용하여 양자화값(QP)을 추정한다(S524).
- <66>        (23) 인터 코딩을 수행한다(S514).
- <67>        (24) 프레임의 수가 N 을 초과한 것으로 판단되면(S516), 역전파(BP) 모델 업데이트 제어를 수행(S526)한다. 생성된 비트들에 따라서 BP는 네트워크를 조정한다. 신경망은 적응적으로 업데이트되고, 뉴우런(neuron)의 전체 갯수는 더이상 변하지 않는다. 계



다가 출력 오차에 따라서 신경망의 구조나 파라미터들을 다음과 같이 변화한다. 비디오 시퀀스의 속성(nature) 때문에, 충분한 정확도를 가지고 출력을 예측하는 것은 용이하지 않다. 특히 장면의 변화(scene change)가 심할 때는 더욱 어렵다.

<68> 이러한 문제점을 해결하기 위하여 네트워크의 업데이트 스킴(scheme)은 두개의 카테고리(category)로 나누어진다. 만일 네트워크를 이용하여 수행한 예측 출력이 실제 출력에 더 근접하면 네트워크의 파라미터들은 온라인 BP에 의해서 업데이트 될 것이다. 만일 오차가 임계치(threshold)를 초과하면 현재의 입력 레코드(record)는 네트워크의 노드(node)를 오버라이트(overwrite)하여 가장 유사한 출력을 얻는다. 다시 말하면, 네트워크의 크기는 변하지 않고 퍼지 기반의 구조(fuzzy base structure)만 변화한다. 상술한 업데이트 스킴의 장점은 네트워크의 크기를 적절하게 조절할 수 있고, 이와 함께 특성변화(variation of the picture characteristics)를 알아낼 수 있다는 것이다.

<69> (25) 포스트 프레임 스킵제어를 수행한다(S508).

<70> (26) 타임 인스턴트 업데이트를 수행한다(S510).

<71> (27) 모든 프레임이 완료될 때까지 상술한 과정을 반복한다.

<72> 도 6은 본 발명의 비트율 제어장치의 블록도이다.

<73> 크게 나누어 프리 인코딩부(610), 인코딩부(620), 포스트 인코딩부(630), 타임 인스턴트 업데이트부(640) 및 판단부(650)로 구성되어 있다.

<74> 프리 인코딩부(610)는 입력된 비디오 프레임을 코딩하기 전에 필요한 여러가지 과정을 수행하며, 초기화 수행부(611), 양자화값(QP) 추정부(612), 조인트 버퍼 제어부(613) 및 타겟 추정부(614)로 구성되어 있다.

- <75> 초기화 수행부(611)는 RD 버퍼에 메모리를 할당하는 등의 기능을 수행한다. 양자화 값(QP) 추정부(612)는 구축된 네트워크를 이용하여 양자화값(QP)을 추정한다. 즉, 구축된 네트워크에 따라서, 인코딩시에 사용될 양자화(QP) 값을 추정한다. 조인트 버퍼 제어부(613)는 네트워크의 구성과 타겟 추정결과에 따라서 RD 버퍼를 수정한다. 타겟 추정부(614)는 RD 버퍼의 상태에 따라서, 어느 정도의 비트들이 프레임 i 에 할당되어야 하는가를 추정한다.
- <76> 인코딩부(620)는 인트라 코딩 수행부(621) 및 인터 코딩 수행부(622)로 구성되어 있어 각각 인트라 코딩과 인터 코딩을 수행한다.
- <77> 포스트 인코딩부(630)는 RD 버퍼 갱신부(631), BP 업데이트 제어부(632), 자가 구축 제어부(632) 및 포스트 프레임 스킵 제어부(634)로 구성되어 있다.
- <78> RD 버퍼는 디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 버퍼로 RD 버퍼 갱신부(631)는 RD 버퍼를 갱신한다. BP 업데이트 제어부(632)는 역전파(BP) 모델 업데이트 제어를 수행(S526)한다. 자가 구축 제어부(632)는 코딩결과에 따라서 네트워크를 구축하고 RD 버퍼를 업데이트 한다. 포스트 프레임 스킵 제어부(634)는 인코딩된 비트 스트림에 따라서 RD 버퍼가 언더플로우나 오버플로우 되지 않도록 다음 프레임의 인코딩 여부를 판단한다.
- <79> 타임 인스턴트 업데이트부(640)는 다음 프레임을 입력받는다.
- <80> 판단부(650)는 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가를 판단한다.

<81> 본 발명은 또한 컴퓨터로 읽을 수 있는 기록매체에 컴퓨터가 읽을 수 있는 코드로서 구현하는 것이 가능하다. 컴퓨터가 읽을 수 있는 기록매체는 컴퓨터 시스템에 의하여 읽혀질 수 있는 데이터가 저장되는 모든 종류의 기록장치를 포함한다. 컴퓨터가 읽을 수 있는 기록매체의 예로는 ROM, RAM, CD-ROM, 자기 테이프, 플로피디스크, 광 데이터 저장장치 등이 있으며, 또한 캐리어 웨이브(예를 들어 인터넷을 통한 전송)의 형태로 구현되는 것도 포함한다. 또한 컴퓨터가 읽을 수 있는 기록매체는 네트워크로 연결된 컴퓨터 시스템에 분산되어, 분산방식으로 컴퓨터가 읽을 수 있는 코드가 저장되고 실행될 수 있다.

<82> 이제까지 본 발명에 대하여 그 바람직한 실시예들을 중심으로 살펴보았다. 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자는 본 발명이 본 발명의 본질적인 특성에서 벗어나지 않는 범위에서 변형된 형태로 구현될 수 있음을 이해할 수 있을 것이다. 그러므로 개시된 실시예들은 한정적인 관점이 아니라 설명적인 관점에서 고려되어야 한다. 본 발명의 범위는 전술한 설명이 아니라 특허청구범위에 나타나 있으며, 그와 동등한 범위 내에 있는 모든 차이점은 본 발명에 포함된 것으로 해석되어야 할 것이다.

#### 【발명의 효과】

<83> 상술한 바와 같이 본 발명은 RD 모델을 자가구축하기 때문에 복수의 샘플, 복수의 양자화기 및 선행 트레이닝과정이 필요하지 않아 비트율 제어에 걸리는 시간이 많이 단축되는 효과가 있다. 그리고, 비트 스트림을 인코딩할 때 RD 모델이 자가 구축되기 때문에, 인코딩된 비트 스트림의 특성을 쉽게 파악할 수 있고 입력된 스트림에 적응적이다. 따라서 RD 모델에 크게 영향을 미치는 샘플 트레이닝 비디오에 의해서 영향받지 않는다.

<84> 그리고, 프레임 기반의 비트스트림을 인코딩할 때, 본 발명은 PSNR(Peak Signal to Noise Ratio) 이 좋고, VM18.0에 정의된 표준방법에 비하여 프레임 스kip이 덜 발생한다.

## 【특허청구범위】

## 【청구항 1】

- (a) 비디오 프레임을 입력받고, 디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 비트율 왜곡(RD) 버퍼에 메모리를 할당하는 초기화 수행 단계;
- (b) 상기 입력받은 비디오 프레임에 대하여 인트라 코딩을 수행하는 단계;
- (c) 상기 RD 버퍼를 업데이트하는 단계;
- (d) 상기 RD 버퍼가 언더플로우나 오버플로우가 되지 않도록 다음 프레임의 인코딩 여부를 판단하여 다음 입력받은 프레임을 스킵하는 포스트 프레임 스킵 수행 단계;
- (e) 다음 비디오 프레임을 입력받아 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가의 여부에 따라서 양자화값을 추정하거나 인터 코딩을 수행하는 단계;
- (f) 상기 인터 코딩을 수행한 후에 상기 입력받은 프레임의 수에 따라서 역전파(back propagation) 모델 업데이트를 수행하거나 자가 구축 제어를 수행한 후에 상기 (d) 단계를 다시 수행하는 단계; 및
- (g) 다음 비디오 프레임을 입력받아 모든 프레임에 대하여 상기 (e) 내지 (f) 단계를 수행하는 단계를 포함하는 것을 특징으로 하는 비트율 제어방법.

## 【청구항 2】

제1항에 있어서, 상기 (e) 단계는

- (e1) 다음 비디오 프레임을 입력받는 단계;
- (e2) 상기 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가를 판단하는 단계; 및

(e3) 상기 판단결과 만일 첫번째 비디오 오브젝트 플레인(VOP)이면 인터 코딩을 수행하고, 그렇지 않으면 상기 RD 버퍼의 상태에 따라서 어느 정도의 비트들이 프레임에 할당되어야 하는가를 추정하는 타겟 추정(target estimation)과, 네트워크의 구성과 상기 타겟 추정결과에 따라서 상기 RD 버퍼를 수정하는 조인트 버퍼 제어를 수행한 후 그 결과를 이용하여 양자화값을 추정하는 단계를 포함하는 것을 특징으로 하는 비트율 제어 방법.

### 【청구항 3】

제2항에 있어서, 상기 (e3) 단계에서 양자화값을 추정하는 것은

(a) 입력 벡터를 입력 레이어에 위치시키는 단계;

(b) 자가구축제어에 의해 구축된 네트워크로부터 예측 출력을 얻는 단계;

(c) 상기 예측출력으로부터 양자화값을 소정의 범위 이내의 오차로 예측하는 단계; 및

(d) 상기 예측한 양자화값을 양자화기에 할당하는 단계를 포함하는 것을 특징으로 하는 비트율 제어방법.

### 【청구항 4】

제3항에 있어서, 상기 (c) 단계는

(c1) 인접 프레임간의 양자화값의 차이가  $\alpha$ 를 넘지 않도록 선택하는 단계; 및

(c2) 상기 예측 출력에 가장 근접한 출력의 양자화값을 찾는 단계를 포함하는 것을 특징으로 하는 비트율 제어방법.

**【청구항 5】**

제1항에 있어서, 상기 (f) 단계는

(f1) 입력받은 프레임의 수가 미리 설정한 소정의 상수값 보다 작은가를 판단하는 단계; 및

(f2) 상기 판단결과, 만일 소정의 상수값 보다 작으면 코딩결과에 따라서 네트워크를 구축하고 상기 RD 버퍼 업데이트 수행하는 자가구축 제어를 수행하고, 그렇지 않으면 역전파(BP) 모델 업데이트 제어를 수행하는 단계를 더 포함하는 것을 특징으로 하는 비트율 제어방법.

**【청구항 6】**

비디오 스트림을 입력받아 코딩에 필요한 버퍼를 초기화하는 프리 인코딩부;

상기 입력받은 비디오 스트림에 대하여 인터 코딩과 인트라 코딩을 수행하는 인코딩부;

상기 코딩된 비디오 데이터에 따라서 상기 버퍼를 업데이트하고, 프레임 스킵제어를 수행하여 비트율을 조절하는 포스트 인코딩부;

다음 프레임을 입력받는 타임 인스턴트 업데이트부; 및

상기 입력받은 프레임이 첫번째 비디오 오브젝트 플레인(VOP)인가를 판단하는 판단부를 포함하는 것을 특징으로 하는 비트율 제어장치.

**【청구항 7】**

제6항에 있어서, 상기 프리 인코딩부는

디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 비트율 왜곡(RD) 버퍼에 메모리를 할당하는 초기화 수행부;

구축된 네트워크를 이용하여 인코딩시에 사용될 양자화값을 추정하는 양자화값 추정부;

상기 RD 버퍼의 상태에 따라서, 어느 정도의 비트들이 입력된 프레임에 할당되어야 하는가를 추정하는 타겟 추정부; 및

상기 구축된 네트워크의 구성과 타겟 추정결과에 따라서 상기 RD 버퍼를 수정하는 조인트 버퍼 제어부를 포함하는 것을 특징으로 하는 비트율 제어장치.

#### 【청구항 8】

제6항에 있어서, 상기 포스트 인코딩부는

디코더로 전송할 인코딩된 비트 스트림을 저장하고 있는 RD 버퍼를 업데이트하는 RD 버퍼 갱신부;

역전파(BP) 모델 업데이트 제어를 수행하는 BP 업데이트 제어부;

코딩결과에 따라서 네트워크를 구축하고 상기 RD 버퍼를 업데이트하는 자가 구축 제어부; 및

인코딩된 비트 스트림에 따라서 상기 RD 버퍼가 언더플로우나 오버플로우 되지 않도록 다음 프레임의 인코딩 여부를 판단하는 포스트 프레임 스킵 제어부를 포함하는 것을 특징으로 하는 비트율 제어장치.

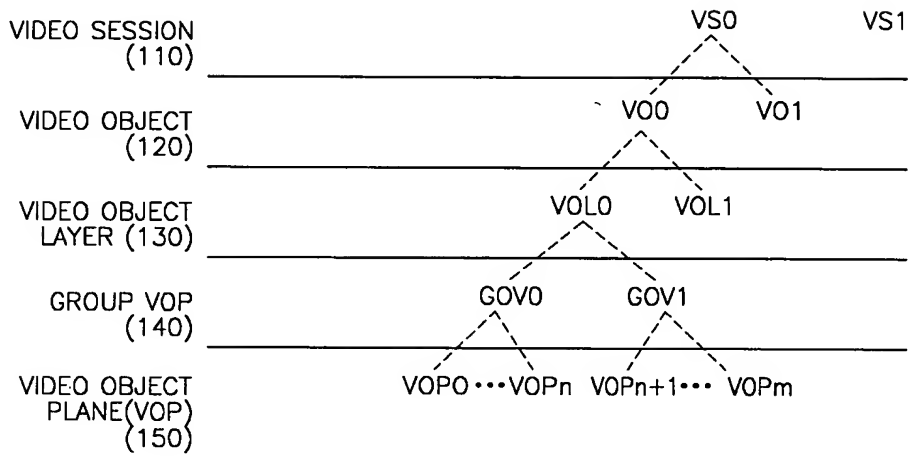


【청구항 9】

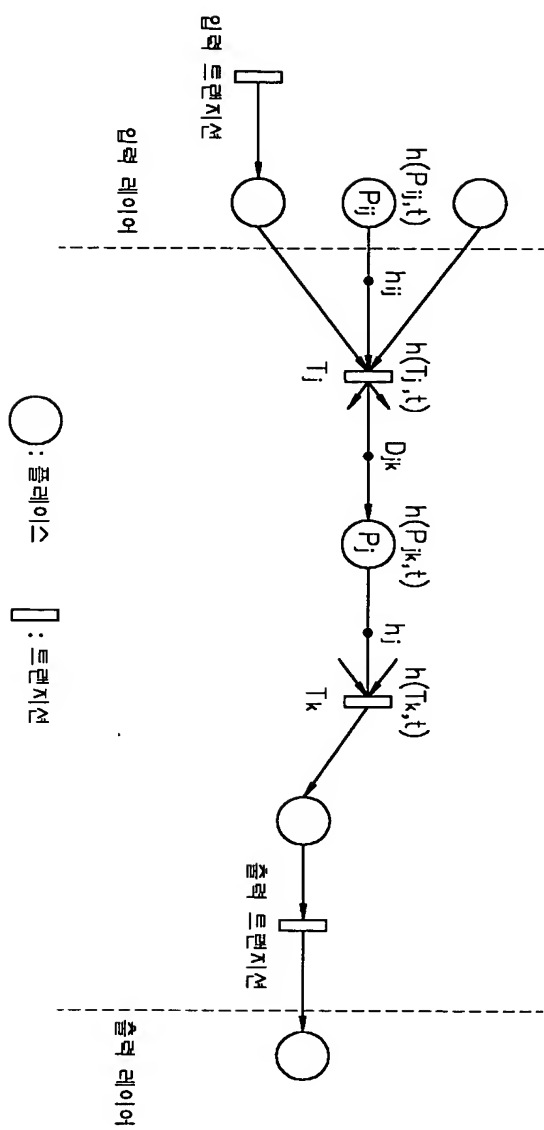
제1항에 기재된 방법을 컴퓨터에서 실행시키기 위한 프로그램을 기록한 컴퓨터로 읽을 수 있는 기록매체.

## 【도면】

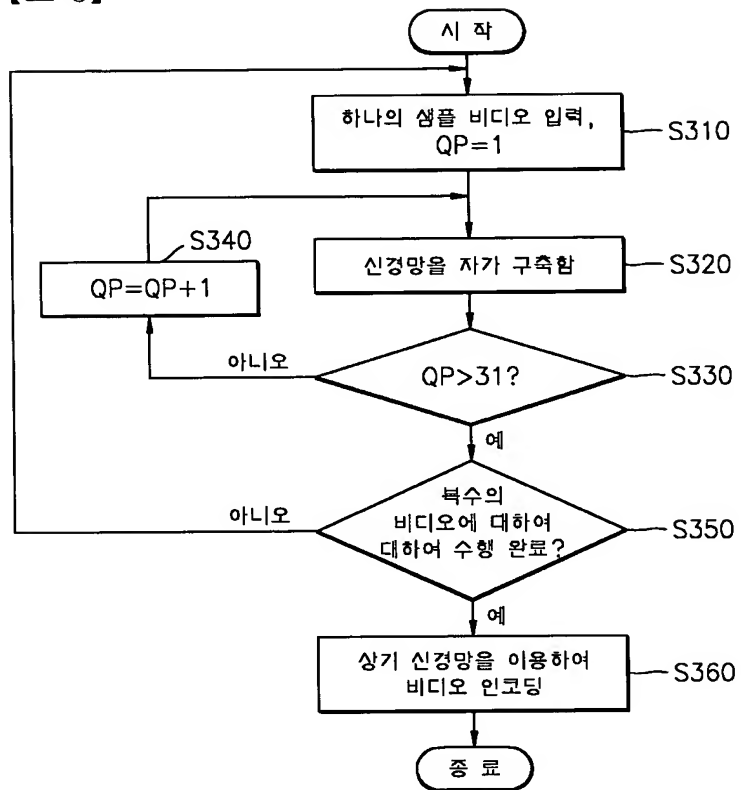
【도 1】



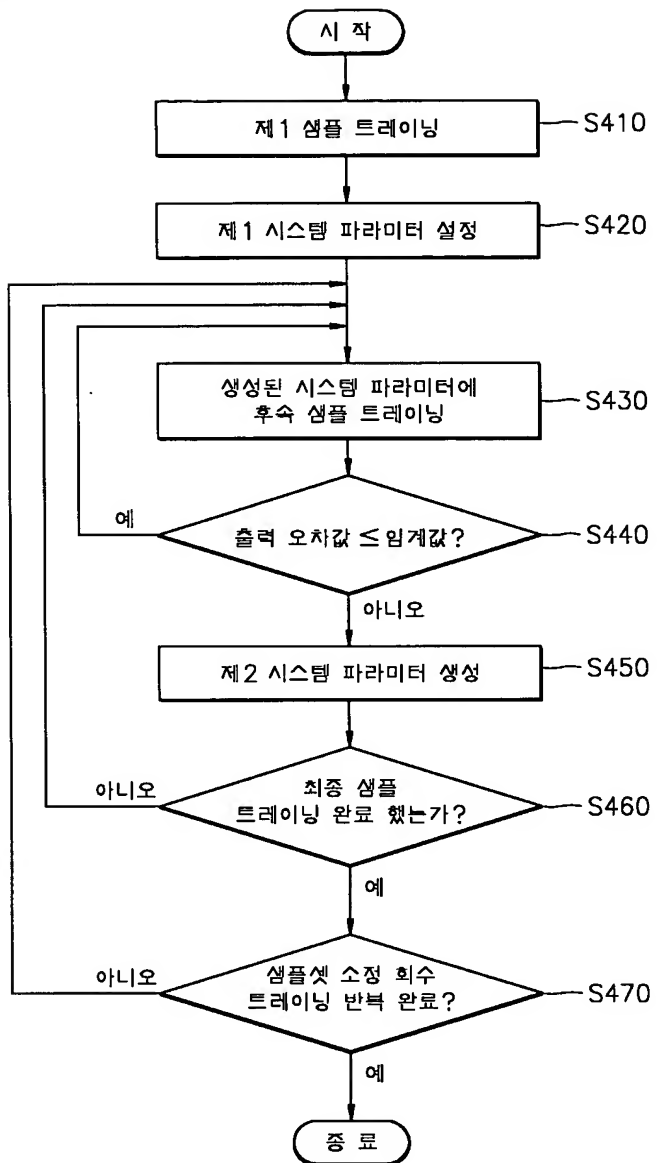
【도 2】



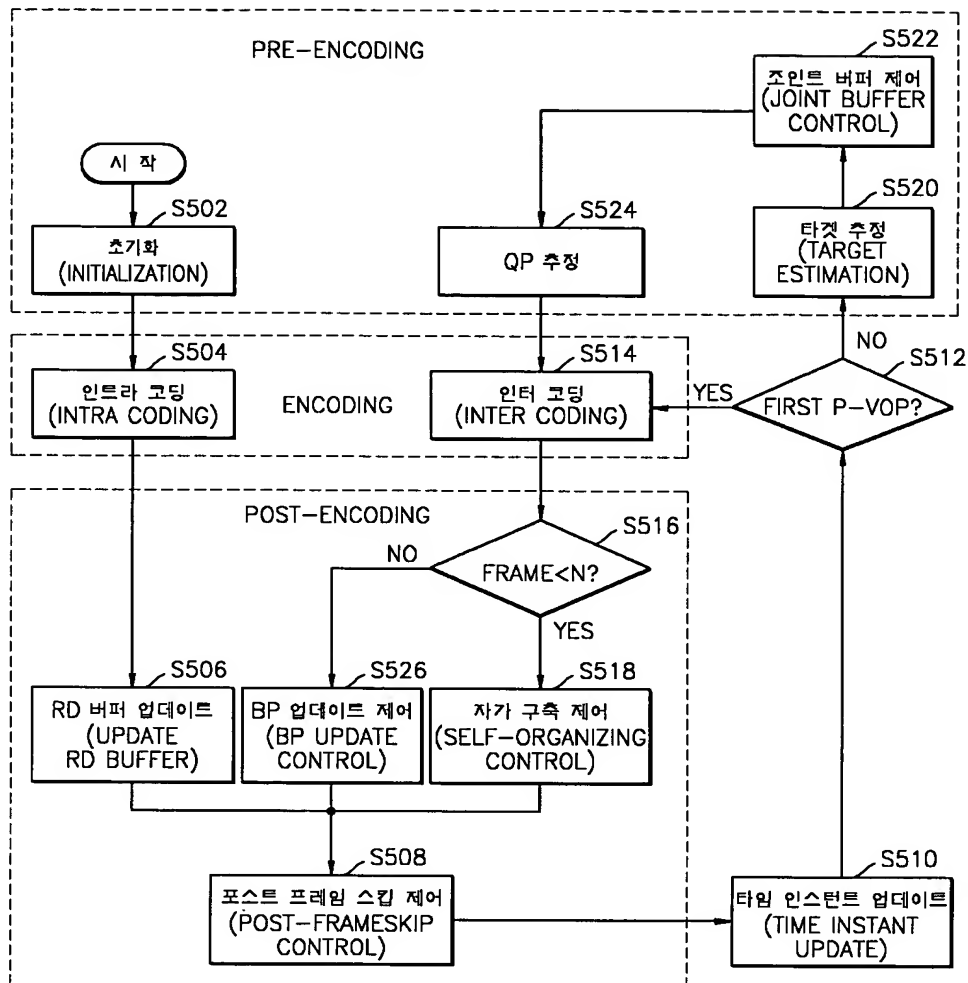
【도 3】



【도 4】



【도 5】



【도 6】

